DOCUMENT RESUME

ED 312 879 FL 018 197

AUTHOR Birdsong, David
TITLE Lux et Verisimilitudo: Judgment Data in SLA Theory.
PUB DATE 89
NOTE 16p.; Paper presented at the Annual Meeting of the Boston University Conference on Language Development (Boston, MA, October 13-15, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Decision Making; *Grammatical Acceptability; *Linguistic Theory; Research Methodology; *Research Problems; Second Language Learning

ABSTRACT

A discussion of the data obtained from grammaticality judgment tasks argues that despite the light shed by these data on problems of second language acquisition theory, there is not yet adequate knowledge of how to interpret those data within a coherent model of performance of the tasks. Therefore, it is concluded, there is no basis for deciding whether the resultant data are plausible and informative or illusory and irrelevant to acquisition theory. It is proposed that judgments of grammaticality be viewed not as behaviors proprietary to the field of linguistics, but as one type of decision-making or judgment-making behavior among many others behaviors. As such, they may be described in terms of procedural or epistemological principles given by more generalized models of cognition and psychophysics. Specifically, performance on grammaticality judgment tasks are examined first within the framework of the Theory of Signal Detectability and then within a more recent framework of concept construction and categorization. Some of the principles of those frameworks are then applied to problems in second language acquisition theory. Finally, some reflections on where second language acquisition research has come from and can go with grammaticality judgment data are offered. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
D. BIRDSONG
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED312879

LUX ET VERISIMILITUDO: Judgment data in SLA theory

B. U. Language Development Conference 1989
David Birdsong, Univ. of Florida

To get the pretentious title out of the way first: The L-word refers to light shed on theoretical problems by empirical data. The V-word refers to the appearance of truth that is the product of our investigations. This "appearance of truth" can be taken in two senses: either as plausibility and credibility, or as illusion, trumpery, and deception. I will argue today that despite the welcome light shed on problems of Second Language Acquisition theory by data from grammaticality judgment tasks, we do not yet know how to interpret these data within a coherent model of performance on such tasks. Thus in the absence of a framework for understanding the behaviors of subjects in judgment tasks, there is no basis for deciding whether the resultant data are plausible and informative or whether they are illusory and irrelevant to acquisition theory.

I realize that what I have just said is beating a dead horse, because for two decades now there's been no end of debates about use and abuse of grammaticality judgments and pleas for better understanding of this type of metalinguistic performance.

The major debaters and pleaders have included: Bever, Carden, Chomsky, Greenbaum, Hill, Labov, Lasnik, Levelt, McCawley, Peters & Ritchie, Ross, and on and on. (See partial listing in references on handout.) However, today, I would like to put debating and pleading behind us, and offer a couple of new ways of looking at grammaticality judgments.

I propose that judgments of grammaticality be viewed not as behaviors proprietary to the field of linguistics, but as one type of decision-making or judgment-making behavior among many others. As such, they may be described in terms of procedural and epistemological principles given by more *generalized* models of cognition and psychophysics. Specifically, I will examine performance on grammaticality judgment tasks first within the framework of the Theory of Signal Detectability introduced by Green and Swets in the 1960's and then within a more recent framework of concept construction and categorization developed by Larry Barsalou. I will apply

Paper presented at the Annual Meeting of the Boston University Conference on Language Development (Boston, MA, October 13-15, 1989).

FL015197

BEST COPY AVAILABLE

some of the principles of these frameworks to problems in Second Language Acquisition theory. Finally, I will offer some reflections on where the SLA research industry has come from and where it can go with grammaticality judgment data.

Judgments of grammaticality, like judgments in general, involve accessing a data base, and using the data as primary input for making a decision. In principle, this data base is linguistic knowledge. The representation of this knowledge is not universally agreed upon, yet it can be safely assumed at least that the data base is stable among adult speakers. Demonstrating this stability empirically is another matter, however. We know from any number of empirical studies (SEE **(1) on handout**) that this presumed knowledge is often distorted or inaccessible across certain task and situational conditions.

Thus Nagata 1988, 1989 finds instability in grammar judgments as a function of contextual variables and variables of induced mental state, Rosenthal 1966, Labov 1975, and Heringer 1970 allude to experimenter effects, Greenbaum 1973, Miller & Isard 1963, and Quirk & Svartvik 1966 document effects of order of presentation, Mohan 1977 refers to a yea-saying factor, Levelt et al 1977 discuss effects of concrete vs. low imagery lexical items.
Still other studies (mentioned in Birdsong, 1989) have shown that this presumed knowledge is also differentially accessible across individuals, as a function of experience and training.

Thus no one can reasonably claim that performance on grammaticality judgment tasks is invariable or that it reflects presumed or idealized linguistic knowledge in any straightforward way. What then does this performance represent? How to interpret the data?

Some promising answers lie in the Theory of Signal Detectability (TSD). The basic premises of TSD are illustrated by a simple anecdotal example. You are listening to the stereo in your house. Under these circumstances, your ability to detect the phone ringing at the opposite end of the house is compromised. Successful detection of the ringing involves distinguishing a signal (the sound of the phone ringing) from the background noise of the stereo. It is with some degree of uncertainty that you render a judgment as to whether the phone is ringing. The Theory of Signal Detectability spells out the factors involved in your determination of whether the phone is ringing.

(SEE (2) on handout)

FACTOR (1) The signal-to-noise ratio. If the stereo us turned up loud, it will be hard to detect the phone's ring. Your facility in separating signal from noise depends in part on the sensitivity of your auditory system.

FACTOR (2) Expectations. If you are expecting a telephone call, it is more probable that you will say the phone is ringing than if you don't believe that a call is a likely event. In most cases; a phone call at four p.m. is more likely than one at 4 a.m. If you've just changed your number and made it unlisted, you will judge the likelihood of the phone's ringing to be low. Internal coding of signal probability is not determined sensorily but cognitively.

FACTOR (3) Motivation. Your tendency to say "Yes, the phone is ringing" or "No, the phone is not ringing" will depend on the relative importance and costs you attach to the event of the phone ringing versus the phone not ringing. For example, you will probably attach more importance to a call you expect from a potential employer than to an expected call from a persistent life insurance salesman. You would be more inclined to say "Yes, the phone is ringing" in the employer scenario than in the insurance salesman scenario. As for the cost you attach to the event of the phone ringing, you might tend to say "No, the phone is not ringing" if you didn't want to interrupt your favorite Bo Diddley song on the stereo, whereas you might say yes if you were concerned about missing a tip for the sixth race at the local horse track. Importance versus cost (or reward versus risk) may be finessed in terms of an interaction of your assessment of the event and your assessment of the consequences of your answer. For example, if you are expecting a call from a potential employer, you may be inclined to say "Yes, the phone is ringing". But if you do not want your spouse to perceive you as being anxious, then there may be a certain risk associated with a "Yes" answer, thus inclining you toward giving a "No" answer.

The Theory of Signal Detectability is able to quantify a value for sensitivity or discriminability (this is referred to as *d prime*). It also permits calculation of response bias, that is, a measure subsuming subjects' ad hoc expectations and motivations. With these measures, plus determinations of signal-to-noise ratios that take into account fixed and conditional

probabilities, TSD is able to describe the conditions under which subjects correctly or incorrectly say "Yes" or "No".

In its broadest conception, the TSD is concerned with the parameters of decision-making and judgment under varying degrees of certainty. According to this theory, judgments are viewed as the product of sensory and cognitive mechanisms. As such, TSD allows for generalization to a number of performance domains besides detection of the telephones ringing.

Indeed, the premises of TSD are compatible with what is generally known about judgments under uncertainty, of which many judgments of grammaticality are a notable subspecies. Let us look first at the applicability of **Factor Two**, subject's expectations. It is commonplace in the literature on judgment under uncertainty to observe biases based on prior expectations. As a trivial example, a run of six sentences judged grammatical can bias the seventh judgment toward ungrammatical, just as a run of six heads in a coin toss influences subjects' guesses for the seventh toss toward tails. This type of behavior, which is based on expectations of the likelihood of distribution of items in a sequence, is documented in Tversky & Kahneman (1974).

Tversky & Kahneman also discuss biases due to the retrievability of instances and to the effectiveness of subjects' search set. In the domain of grammaticality judgments, biases may be a function of subjects' ability to retrieve examples of similar sentences from memory, and of their ability to construct or imagine contexts in which items to be judged might occur. Contributing to such biased expectations may be recency effects, as well as the salience, frequency, and imagery of lexical items in stimulus sentences.

As for **Factor Three**, the notoriously slippery and idiosyncratic variable of motivation, a few simplistic illustrations from the domain of L2 acquisition should suggest its role in grammaticality judgments. It is easy to imagine, for example, how respondents' learning experiences and learning goals could prime differential motivations toward yes and no responses. Presumably, learners who take a highly analytical approach to L2 learning, or those who place a premium on accuracy, may as a result valorize their ability to detect structural deviance. In such a case the rewards of detecting ungrammatical items would be high, as would the perceived *risks* of *not* detecting such items. Such a motivation could bias subjects toward giving "No" answers. One can imagine another scenario whereby more advanced subjects

familiar with target language variability would associate liberality of judgment with native speakerhood; in such a case, a "Yes" for a quite marginal sentence would carry a high reward, a "No" for such a sentence might carry a risk. This behavior can be related to the finding of Zobl (this conference) that multilinguals tend to accept marked structures in L2. It is also similar to the "going out on a limb" phenomenon, noted extensively in the SLA literature, whereby given the possibility of scalar responses beginners tend to cluster their responses about intermediate values, while more advanced subjects' answers tend to be more extreme. Such varieties of self-consciousness are reminiscent of the effects of self-awareness in judgments found by Nagata 1989, who follows on the work of Carroll, Bever & Pollack 1981.

Let us return now to **Factor One**, Signal-to-Noise ratios. For the Theory of Signal Detectability to apply plausibly to grammaticality judgments, one has to extend the notion of signal-to-noise ratios beyond the domain of primitive sensory systems to embrace cognitively-based perceptual processes. Under such a view, noise may be understood as the ambient stream of well-formed utterances. Against this background of noise, occasional signals in the form of ungrammatical utterances are produced. As in the standard model of signal-to-noise ratios, the perceptual variable is one of sensitivity, in this case sensitivity to ill-formedness. Consistent with the literature on the subject, this sensitivity can be viewed rather straightforwardly as a matter of training and experience, with normal distributions of individual differences. That is, some people are known to be more adept than others at detecting deviance in normal speech. Presumably, this skill transfers to more formal experimental situations. For example, one might anticipate skill effects in timed grammaticality judgment tasks, where subjects' baseline, ecologically-established sensitivity to deviance is a latent independent variable. The introduction of perceptual variables in the rendering of linguistic intuitions is consonant with work by Gerken & Bever (1986), among others, who argue for an interaction of perceptual variables and mental representations of linguistic universals.

I offer the Theory of Signal Detectability as a framework for examining procedural aspects of judging grammaticality because it incorporates into a single model both perceptual variables and subjects' biases. I believe that the model could be enhanced by further elaboration of the heuristics or rules of

thumb which inform subjects' expectations and biases and which guide judgments (see, e.g., Fishhoff, 1988; Kahneman, Slovic & Tversky, 1982). One widely-cited heuristic is the *availability heuristic*. Events that occur more frequently than others are more available in memory, and are conferred face validity. Other factors affecting availability include salience and recency effects. Another heuristic is *representativeness*. Subjects tend to make judgments based on how well items represent salient features normally associated with targets. Kahneman & Tversky (1972) showed that this heuristic can override other factors such as prior probability. Clearly, the interactions of the representativeness heuristic and probabilistic factors are pertinent to the process of making judgments of grammaticality, and especially in cases where subjects assume (or are told by the experimenter) that there are equal numbers of grammatical and ungrammatical items.

It is with respect to the notion of representativeness that Barsalou's theory of concepts and categorization offers ways of understanding the epistemological factors relative to decision making in grammaticality judgment tasks.

Barsalou (1987) makes a critical distinction between category and concept. Categories are cognitive structures which may have finite or infinite membership. Some categories may be formal (e.g., ODD NUMBERS, SQUARES), while others may be goal-derived (e.g., THINGS TO TAKE ON A CAMPING TRIP). Still others may not be established in memory, and indeed are rarely if ever thought about. Barsalou has shown empirically, for example, that people are able to create and manipulate such categories as WAYS TO ESCAPE BEING KILLED BY THE MAFIA and THINGS THAT COULD FALL ON YOUR HEAD. Categories typically display gradedness or prototype effects, such that in experimental performance some members of a category are perceived as better exemplars of that category than others. Barsalou, like Lakoff and Rosch, insists that the graded structure of categories is a behavioral effect, and not an epistemological given.

In this scheme, a concept is invoked or constructed to index (or, in Barsalou's terminology, *represent*) a category. Thus, having wings is a concept that represents the category of BIRD. Barsalou rejects the classical

association of concepts with defining properties, distinctive features, or criteria for membership. Instead, the term concept refers to particular information used to represent a category on a particular occasion. That is, concepts are not necessarily invariant or stable. More precisely, the concept contains information that provides relevant expectations about the category in a given context as well as information about that category in most contexts. [examples to follow] Among Barsalou's reasons for describing concept in this way is the fact that defining properties for categories are often not available. Thus Lakoff's famous example of the category of MOTHER. The concept "birth-giver" doesn't always work, because there are adoptive and foster mothers. Even "female" fails as an invariant concept since there are females who have given birth and have since had a sex-change operation.

And even when criterial definitions for categories do exist, Barsalou argues on the basis of empirical evidence that such definitions do not always operate in all people's representations of categories. To illustrate his point, Barsalou cites the category ANIMAL. Pet owners may generate a concept for animals as small and domesticated. Hunters on the other hand might use a concept of animals as being large and/or wild. Thus, a house cat may be judged more animal-like by pet owners, but as a poorer exemplar of the category ANIMAL by hunters.

The variability of concepts is also documented in a variety of ad hoc and context-dependent behaviors. For example, Barsalou notes that the concept "floating" is not normally associated with the category BASKETBALL. However, when subjects are told that someone in a boating accident used a basketball as a life preserver, the concept "floating" becomes activated as a concept.

An application of Barsalou's framework is given in 4. Here, the so-called invariant concept which presumably represents this category SQUARE is "plane figure with 4 sides of even length joined at right angles." But when is a square not a square? In certain contexts, such as 4a, squares are not perceived and judged to be squares. The same may be true in other contexts as well, as in 4b. Here two triangles are being placed together. The resultant figure could be, by definition, a square -- or -- two triangles. For it to be judged a square, an individual would have to invoke, or indeed, <u>create</u>, a variable concept, namely, "two congruent right triangles isosceles joined along the length of their hypotenuses". Further manipulation of contextual

features may result in considerable variability across subjects and instability within subjects. Imagine for example that the triangles were originally separated and rotated, and then manually joined to yield the square, as in 4c on the handout. In such a case subjects might be primed to perceive two triangles, and thus be less inclined to judge the figure a square than under a condition where the triangles were already joined. [note parallel to diachronic and synchronic notions of grammaticality in language] And similarly, referring to **4d** and **4e**, the perceived squareness of the figure may vary, depending on whether the diagonal line representing the triangles' hypotenuses is prominent (**4d**)--as opposed to a condition whereby the diagonal is obscured or is not present at all (**4e**).

With these metacognitive behaviors, there are abundant parallels in judgments of sentence grammaticality. At one time or another, we have probably all concocted shaggy-dog stories to establish contexts for making starred sentences seem OK. As McCawley (1985) points out, linguists who offer their intuitions often are not grappling so much with questions of grammaticality but rather reporting their success in imagining a context where the sentence in question would sound OK. I believe the same can be said for most normals as well. Such behavior would be consistent not only with the Barsalou framework, but also with the TSD framework, where imaginability and effectiveness of search routines are variables (see Factor 2).

This could be one source of the attested instability of grammaticality judgments. Another source could be *ad hoc* concepts. As a common example, an individual may reject sentences like (**5**), with stranded prepositions and improper case marking, on one occasion and accept them on the next, depending on how prescriptivist the person's *ad hoc* concepts which represent the category of well-formedness.

In (**6**) I suggest how grammaticality judgments might be accommodated within Barsalou's framework. First one must assume a category of WELL-FORMEDNESS. The *invariant concepts* that represent this category could be derived from principles, parameters, constraints, phrase structure rules, etc. As suggested by copious evidence in the literature, the *variable concepts* might include parsability, euphony, and the sentence's status as semantically

non-anomalous. These variable concepts may be invoked or created as a function of contextual factors or situational factors.

The operation of variable concepts are charmingly illustrated in Boutet's (1986) study of 6-to-11 year-old native French speaking children. The instructions given to these subjects were "Dis-moi pour toi ça fait une phrase" ('Tell me [if] for you this makes [i.e., is] a sentence'). Curiously, nearly 30% of Boutet's subjects rejected the item in (7). *Quand ta grand-mère arrivera-t-elle?* Those subjects' explanation in post-hoc think alouds was that the item is a question, and therefore not a sentence. The generation of such a concept to represent well-formedness is not so surprising, given that this was the only item in the corpus of stimuli that was an interrogative. Thus, consistent with Barsalou's framework, we see the effects of priming or prior context (here, the types of items in the experimental set).

In Boutet's study, several other concepts appeared to have been created or invoked on an ad hoc basis, e.g., "the words make a sentence of a good length", "sounds OK if written in a telegram", "needs a comma". These concepts varied both across and within subjects. These concepts may be compared with those often invoked by adults when judging grammaticality, e.g., those suggested in (6).

Parenthetically, I might mention that the creation of such variable concepts is discussed with a much broader scope of application in the literature on folk taxonomies. (For the relevance of folk taxonomies to grammaticality judgments, see Odlin, 1986). I also think variable concepts can be accommodated within the expanded TSD framework suggested earlier. I have in mind the idea that the *availability* of a concept must be taken into account, thus allowing one to describe a condition whereby certain variable concepts are more available under immediate demands of the task than are other variable concepts, and are also more available than are presumed invariant concepts.

Now that the basics of Barsalou's work and TSD have been sketched, what is the relevance of these ideas to the future of L2 acq. theory? I think the answer to this question must be situated within the current state of SLA theoretical inquiry. Some L2 acq researchers hold to the naive assumption

that subjects home in on relevant grammatical structure and base their judgments solely on knowledge of that structure. Others regard judgment data with outright scorn. In the face of such polarized opinion, use of grammaticality judgments persists, as attested by the fact that at least half of the empirical papers at this year's LARS symposium in Utrecht cited grammaticality judgment data. Often the interpretations of grammaticality judgment data in the SLA literature are no more than exercices in mind-reading--that is, unprincipled post-experimental speculations, with imputations to learners of assorted strategies, principles, mini-grammars, and the like. Those of you who are familiar with the literature know what I'm talking about.

I believe that the field of SLA research is justifiably regarded by outsiders (and by some insiders as well) as churning about amid a lot of data it doesn't really know what to do with. It is therefore not surprising that SLA research does not fit very well the profile of a scientific specialty as sketched by De Mey (1982) (skip to # **11** on handout). Where we depart most from this schema is in the row marked "methodological orientation". Even excluding the problems posed by naturalistic L2 data and elicited imitation data, (on the latter, see new paper by Craig Chaudron), it doesn't seem as if we have come close to stage 3 with respect to methodological orientation. (I dare say with respect to "cognitive context" that many of us skipped right over the first three stages and went straight to stage 4.)

Where then do we go from here? First, let's state the obvious. The Barsalou and TSD frameworks reaffirm two simple principles of grammaticality judgments. (SEE **(8)**)

ONE: Different individuals can generate the same judgment but for different reasons.

TWO: It is conceivable that two individuals could possess identically represented grammars, but could on either an ad hoc or systematic basis generate different grammaticality judgments.

These principles add up to the conclusion that judgments are not necessarily deterministic with respect to a presumed stable base of knowledge. But this is to beat the dead horse again, so let's turn the issue around. Let's put

aside the momentarily the question of whether we can learn anything about linguistic knowledge from grammar judgments. Let's consider instead the question of whether we can learn anything about the decision-making processes that mediate grammatical knowledge and grammaticality judgments.

This idea can be fleshed out with a concrete application. Coppieters (1987) examined native speakers' and near-natives' judgments of French sentences and found what he called "competence differences" between the two groups. While a few of the items exemplified core grammatical distinctions, most of the items involved peripheral structures and subtle distinctions on the border between grammar and stylistics. These fringe sentences were chosen deliberately in hopes that the judgments would require deliberation and not automatic responses. Coppieters' think-aloud protocols revealed numerous inter-group differences with respect to confidence in their judgments, their criteria for judgment, and their homespun linguistic analyses of the items they judged ungrammatical. Faced with such divergent data, Coppieters reasserts his conclusion of competence differences between natives and near-natives. But what the divergent think-aloud data suggest to me are profound performance differences--differences that have yet to be investigated within a coherent framework of judgment under varying degrees of certainty. For example, what are the intergroup differences and similarities in terms of sensitivity to ungrammaticality? Does sensitivity vary across native language and target language stimuli? What kinds of expectations and biases do the two groups manifest? Are the rewards and risks for detecting linguistic deviance the same for the two groups? Are the variable concepts of the two groups the same? Do the groups respond similarly to factors of recency, salience, and context? Imagine an abstraction of the answers to these questions in terms of distributions of points about some measures of central tendency: Do the distribution curves of natives and non-natives overlap? Are they congruent or are they different?

In stating the issue in this way, I am picking up the gauntlet laid down by Lasnik (1981), who insists that judgments must be regarded as performance phenomena (SEE (9)). My first response to this challenge is to suggest that judgment behaviors are worthy of investigation in their own right. The basic question is this: How do people think under the conditions of grammaticality judgment tasks; what are the components of the decision-making process?.

For SLA research, it is reasonable to ask whether learners and native speakers think the same way under such conditions. Of course, the question can be extended to different groups of non-natives as well. I feel that such a line of inquiry can be included under the umbrella of the current research emphasis on cross-cultural differences in problem-solving and in factors contributing to intelligence.

My second suggestion is that once we've compared the decision-making processes of natives and learners (and / or learners from different linguistic backgrounds) within a framework of judgment under varying degrees of certainty, then we may be in a position to probe questions of linguistic knowledge. As a point of departure for determining linguistic knowledge via grammaticality judgments, we must consider various logical possibilities. The oversimplifed permutations are given in **(10)** on the handout.

I.	convergent decision making processes	convergent gramm. judgments
II.	convergent decision making processes	divergent gramm. judgments
III.	divergent decision making processes	convergent gramm. judgments
IV.	divergent decision making processes	divergent gramm. judgments

As I suggested earlier, researchers have tended to assume only permutations I and II, and to ignore possibilities III and IV. Clearly, theories of L2 linguistic competence based on judgment data cannot afford to sweep these under the rug.

In summary, what I've been talking about today are new ways of looking at judgment data. These perspectives offer some hope of helping us eventually to distinguish between the illusion of truth and the real thing.

LUX ET VERISIMILITUDO: Judgment data in SLA theory
Boston University Language Development Conference, 1989
David Birdsong
University of Florida

1. Nagata (1988, 1989): Instability as function of contextual variables and variables of induced mental state
   Rosenthal (1966), Labov (1975), Heringer (1970): Experimenter effects
   Greenbaum (1973), Miller & Isard (1963), Quirk & Svartvik (1966): Order of presentation effects
   Mohan (1977): Yea-saying factor
   Levelt et al (1977): Effects of concrete vs. low imagery lexical items
   Birdsong (1989): Effects of experience and training

2. Theory of Signal Detectability (TSD): Green & Swets (1966)

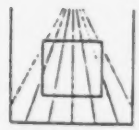
   Factor 1: Signal-to-noise ratios
   Factor 2: Expectations
   Factor 3: Motivation

3. Barsalou (1987): Concepts are invoked (or created) to represent categories

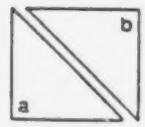
4.

CATEGORY	SQUARE
"Invariant" concept	Plane figure with 4 sides of even length joined at right angles
variable concept	2 congruent right isosceles triangles joined along the length of their hypotenuses

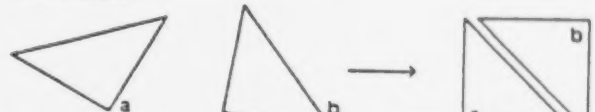
4a.



4b.



4c.



4d.

4e.

BEST COPY AVAILABLE

5. Who are you talking with?

6.

CATEGORY	WELL-FORMEDNESS
"Invariant" concepts	derived from principles, parameters, constraints, ps rules, etc.
variable concepts	parsability, euphony, semantically OK, etc.

7. Quand ta grand-mère arrivera-t-elle? (from Boutet, 1986)

8. Different individuals can generate the same judgment but for different reasons.
Identical grammars can yield divergent grammaticality judgments.

9. Lasnik (1981: 20): "Grammaticality judgments are often incorrectly considered as direct reflections of competence. ... responding to a grammaticality query is an instance of *performance*.

10.

I. convergent decision making	convergent grammaticality judgments
II. convergent decision making	divergent grammaticality judgments
III. divergent decision making	convergent grammaticality judgments
IV. divergent decision making	divergent grammaticality judgments

11. (from De Mey, 1982)

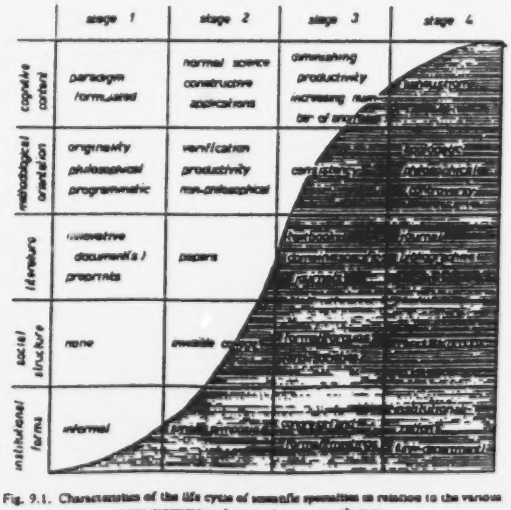


Fig. 9.1. Characteristics of the life cycle of scientific specialties in relation to the various stages superimposed upon a logistic growth curve.

Barsalou, L. W. (1987). The instability of graded structures: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). Cambridge: Cambridge University Press.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence.* Berlin & New York: Springer-Verlag.

Bley-Vroman, R. W., Felix, S. W., & Ioup, G. I. (1988). The accessibility of universal grammar in adult language learning. *Second Language Research, 4,* 1-32.

Boutet, J. (1986). Jugements d'acceptabilité et raisonnements métalinguistiques. *Etudes de Linguistique Appliquée, 62,* 40-58.

Carden, G. (1970). Discussion of J. T. Heringer, "Research on quantifier-negative idiolects". In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society* (p. 296). Chicago: Chicago Linguistic Society.

Carroll, J. M., Bever, T. G., & Pollack, C. R. (1981). The non-uniqueness of linguistic intuitions. *Language, 57,* 368-382.

Chaudron, C. (in preparation). Review of elicited imitation data.

Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language, 63,* 544-573.

de Mey, M. (1982). *The cognitive paradigm.* Dordrecht: Reidel.

Fischhoff, B. (1988). Judgment and decision making. In R. J. Sternberg & E. E. Smith, (Eds.), *The psychology of human thought.* Cambridge & New York: Cambridge University Press.

Gerken, L., & Bever, T. G. (1986). Linguistic intuitions are the result of interactions between perceptual proceses and linguistic universals. *Cognitive Science, 10,* 457-476.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge & London: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cognitive Psychology, 2,* 430-454.

Labov, W. (1975). *What is a linguistic fact?* Lisse: Peter de Ridder.

Lasnik, H. (1981). Learnability, restrictiveness, and the evaluation metric. In C. L. Baker & J. J. McCarty, (Eds.), *The logical problem of language acquisition* (pp. 1-29). Cambridge, MA: MIT Press.

Levelt, W. J. M. (1974). *Formal grammars in linguistics and psycholinguistics. Volume III. Psycholinguistic applications.* The Hague: Mouton.

McCawley, J. D. (1985). [Review of *Grammatical theory: Its limits and possibilities*]. *Language, 61,* 668-679.

Nagata, H. (1988). The relativity of linguistic intuitions: The effect of repetition on grammaticality judgments. *Journal of Psycholinguistic Research, 17,* 1-17.

Nagata, H. (1989). Effect of repetition on grammaticality judgments under objective and subjective self-awareness conditions. *Journal of Psycholinguistic Research, 18,* 255-269.

Odlin, T. (1986). On the nature and use of explicit knowledge. *International Review of Applied Linguistics, 24,* 123-144.

Peters, P. S., & Ritchie, R. (1973). On the generative power of transformational grammars. *Information Sciences, 6,* 49-83.

Ross, J. R. (1979). Where's English? In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 127-163). New York: Academic Press.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185,* 1124-1131.